

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

Ford, Vanessa

From: Ford, Vanessa
Sent: Friday, June 07, 2002 8:53 AM
To: STIC-Biotech/ChemLib
Subject: In Re: 10054647 Sequence Search

Please search SEQ ID NO:2.

Please include an interference search.

Please disregard the previous email about this case.

Biotechnology Patent Examiner
Office: CM1 8D17
Mailbox: CM1 8E12
Phone: 703.308.4735

GenCore version 4.5
Copyright (C) 1993-2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2002, 14:20:58 ; Search time 30.07 seconds
(without alignments)
73.877 Million cell updates/sec

Title: US-10-054-647-2

Perfect score: 110

Sequence: 1 NNTTGVFLKQDWGATIKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	288	20	AAU06959
2	96	87.3	280	20	AAU06948
3	90	81.8	280	19	AAW51094
4	90	81.8	280	21	AAW51094
5	90	81.8	280	22	AAU04198
6	82	74.5	256	20	AAU06942
7	82	74.5	276	19	AAW51095
8	82	74.5	276	21	AAW51095
9	82	74.5	276	22	AAU04199
10	82	74.5	280	19	AAW51089
11	82	74.5	280	21	AAW51089

Major antigenic pr
E. chafeensis omp-
E. canis p30-2 pro
Ehrlichia canis im
E. canis p30-4 pro
Ehrlichia chaffeen
E. chafeensis omp-
Ehrlichia chaffeen
Variable surface a
Ehrlichia canis im
E. canis p30-1 pro
Cowdria ruminantiu
Major antigenic pr
Ehrlichia chaffeen
E. chafeensis omp-
Ehrlichia chaffeen
Variable surface a
E. chafeensis omp-
Variable surface a
Ehrlichia canis VS
Ehrlichia canis pa
Ehrlichia canis im
Ehrlichia canis im
Propionibacterium
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
C glutamicum prote
Human EST encoded

ALIGNMENTS

RESULT 1
AAU06959
ID AAU06959 standard; Protein; 288 AA.
AC AAU06959;
XX
XX
DT 05-JUL-1999 (first entry)
XX
DE E. canis p30 protein.
XX
KW Outer membrane protein; omp; Ehrlichia chafeensis; E. canis; p30;
KW detection; dog.
XX
OS Ehrlichia canis.
XX
PN MO9913720-A1.
XX
PD 25-MAR-1999.
XX
PF 18-SEP-1998; 98WO-US19600.
XX
PR 19-SEP-1997; 97US-0059353.
XX
XX (OHIS) UNIV OHIO STATE.
XX
XX Ohashi N, Rikihisa Y;
XX
DR WPI; 1999-254290/21.
DR N-PSDB; AAX34759.
XX
XX Novel outer membrane proteins from Ehrlichia chafeensis and
XX Ehrlichia canis
XX
PS Disclosure: Fig 19B; 55pp; English.

XX The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 XX SQ Sequence 288 AA;

Query Match 100.0%; Score 110; DB 20; Length 288;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGTGVLKQDWDGATIKD 20
 |||||
 Db 60 ntgtgvlkqdwgdgatkid 79

RESULT 2

AAV06948
 ID AAY06948 standard; Protein; 280 AA.

XX AC AAY05948;

XX DT 05-JUL-1999 (first entry)

XX DE E. chaffeensis OMP-1F protein.

XX KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 XX detection; dog.

XX OS Ehrlichia chaffeensis.

XX PN WO9913720-A1.

XX PD 25-MAR-1999.

XX PF 18-SEP-1998; 98WO-US19600.

XX PR 19-SEP-1997; 97US-0059353.

XX PA (OHIS) UNIV OHIO STATE.

XX PI Ohashi N, Rikihisa Y;

XX PR WPI; 1999-254290/21.

XX DR N-PSDB; AAX34748.

XX PT Novel outer membrane proteins from Ehrlichia chaffeensis and
 XX Ehrlichia canis

XX PS Claim 16; Fig 8B; 55pp; English.

XX The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 XX SQ Sequence 280 AA;

Query Match 87.3%; Score 96; DB 20; Length 280;
 Best Local Similarity 94.4%; Pred. No. 2.5e-07;
 Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGTGVLKQDWDGATI 18
 |||||
 Db 60 ntgtgvlkqdwgdgstl 77

RESULT 3

AAW51094
 ID AAW51094 standard; Protein; 280 AA.

XX AC AAW51094;

XX DT 14-SEP-1998 (first entry)

XX DE Ehrlichia chaffeensis VSA4 protein.

XX KW MAP1 homologue; variable surface antigen; VSA4; rickettsia;
 XX DNA vaccine.

XX OS Ehrlichia chaffeensis.

XX FH Key Location/Qualifiers
 XX Peptide 1..25
 XX /note= "putative signal peptide"

XX PN WO9816554-A1.

XX PD 23-APR-1998.

XX PF 17-OCT-1997; 97WO-US19044.

XX PR 17-OCT-1996; 96US-0733230.

XX PA (UYFL) UNIV FLORIDA.

XX PI Barbet AF, Burridge MJ, Santa RR, Mahan SM, McGuire TC;
 XX Nyika A, Rurangirwa FR;

XX DR WPI; 1998-251232/22.

XX DR N-PSDB; AAV071179.

XX Composition containing nucleic acid encoding rickettsial antigen -
 XX useful for, e.g. stimulating protective immune response in humans or
 XX animals

XX PS Claim 3; Fig 2B; 39pp; English.

XX This is the full-length variable surface antigen VSA4 protein of
 CC Ehrlichia chaffeensis. Its amino acid sequence was deduced from a
 CC partial open reading frame (ORF4) of a genomic locus (see AAV071179)
 CC of E. chaffeensis that was obtained on the basis of homology to the
 CC major antigenic protein MAP1 (see AAW51088) of Cowdria ruminantium.
 CC This genomic locus included 5 ORFs encoding similar, but
 CC non-identical proteins (see AAW51091-95). A claimed composition
 CC comprises a nucleic acid (see AAV07176-82) encoding a polypeptide
 CC (see AAW51088-99) that elicits a protective immune response against a
 CC rickettsial pathogen. The nucleic acid is used, in human or
 CC veterinary medicine, in vaccines to protect against Rickettsia,
 CC Ehrlichia, Anaplasma and Cowdria species. The Ehrlichia antigenic
 CC polypeptides can be used diagnostically to detect antibodies
 CC associated with Ehrlichia infection (claimed).
 XX
 XX SQ Sequence 280 AA;

Query Match 81.8%; Score 90; DB 19; Length 280;
 Best Local Similarity 88.9%; Pred. No. 2.3e-06;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGTGVLKQDWDGATI 18
 |||||
 Db 60 ntgtgvlkqdwgdgstl 77

RESULT 4

AAB36188
 ID AAB36188 standard; Protein; 280 AA.

XX

```

AC AAB36188;
XX
DT 02-MAR-2001 (first entry)
XX
DE Ehrlichia chaffeensis partial VSA4.
XX
KW Ehrlichia chaffeensis; VSA4; variable surface antigen 4; MAP1;
KW major antigenic protein 1; antirickettsial; vaccine; gene therapy;
KW Rickettsia; Cowdria; Anaplasma; map2; lhworf3; 4hworf1; 18hworf1;
KW 3gdorf3.
XX
OS Ehrlichia chaffeensis.
XX
PN WO2000065063-A2.
XX
PD 02-NOV-2000.
XX
PF 21-APR-2000; 2000WO-US10886.
XX
PR 22-APR-1999; 99US-0130725.
XX
PA (UYFL ) UNIV FLORIDA.
XX
PI Barbet AF, Bowie MV, Ganta RR, Burridge MJ, Mahan SM, McGuire TC;
PI Rurangirwa FR, Moreland AL, Simbi BH, Whitnair WW, Allemen AR;
XX
DR WPI; 2000-679675/66.
DR N-PSDB; AAC68705.
XX
PT New polynucleotides useful as DNA vaccines for conferring immunity to
PT rickettsial infection e.g. heartwater caused by Cowdria ruminantium,
PT comprises major antigenic protein 1 or 2 gene of rickettsial pathogens
XX
PS Claim 3; Page 45-46; 63pp; English.
XX
CC The present sequence shows a high degree of similarity to the major
CC antigenic protein 1 (MAP1) of Ehrlichia sp. The MAP1 gene may be
CC used in a vaccines to protect animals or humans against rickettsial
CC diseases caused by a organisms of Rickettsia sp., Ehrlichia sp.,
CC Anaplasma sp. or Cowdria sp. The vaccine elicits an immune response
CC protective against the rickettsial pathogen. The nucleic acid vaccines
CC can be driven by the human cytomegalovirus (HCMV) enhancer-promoter.
CC Cowdria ruminantium genes designated map 2, lhworf3, 4hworf1, 18hworf1
CC and 3gdorf3 may be used in therapeutic and diagnostic applications. The
CC polypeptides are useful for detecting antibodies associated with
CC infection by a rickettsial pathogen whilst the polynucleotides may be
CC used to detect the presence of rickettsial nucleic acids.
XX
SQ Sequence 280 AA;

Query Match 81.8%; Score 90; DB 21; Length 280;
Best Local Similarity 88.9%; Pred. No. 2.3e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNTTGVFLKQDWGATI 18
   ||| ||||| ||||| |||
DB 60 nttagvfglkqdwgstl 77

RESULT 5
AAU04198
ID AAU04198 standard; Protein; 280 AA.
XX
AC AAU04198;
XX
DT 23-OCT-2001 (first entry)
XX
DE Variable surface antigen 4 (VSA4) from Ehrlichia chaffeensis.
XX
KW Major antigenic protein; MAP; vaccine; immunogenic; rickettsia;
KW infection; heartwater; diagnostic; variable surface antigen; VSA.

```

```

XX OS Ehrlichia chaffeensis.
XX
PN US6251872-B1.
XX
PD 26-JUN-2001.
XX
PF 17-OCT-1997; 97US-0953326.
XX
PR 17-OCT-1996; 96US-0733230.
XX
PA (UYFL ) UNIV FLORIDA.
XX
PI Barbet AF, Ganta RR, McGuire TC, Burridge MJ, Nyika A;
PI Rurangirwa FR, Mahan SM, Bowie MV, Allemen AR;
XX
DR WPI; 2001-424487/45.
DR N-PSDB; AAS07578.
XX
PT New MAP2 genes and polypeptides useful as vaccines for conferring
PT immunity to human and animal rickettsial diseases, e.g. heartwater, or
PT as molecular markers in nucleic acid analysis procedures
XX
PS Example 3; Fig 2A-2B; 30pp; English.
XX
CC The sequence represents the amino acid sequence of variable surface
CC antigen 4 (VSA4) isolated from Ehrlichia chaffeensis, which
CC has similarity to major antigen protein (MAP). The MAP polynucleotides
CC and polypeptides are useful as vaccines for conferring immunity to
CC rickettsia infection, including Cowdria ruminantium causing heartwater.
CC The MAP polynucleotides may be used as molecular markers in nucleic acid
CC analysis procedures, and to produce the MAP polypeptides, which may
CC be used to raise antibodies that are reactive with the polypeptides.
CC The nucleic acids may further be used as probes to identify
CC complementary sequences within other nucleic acid molecules or genomes,
CC where such probes can be applied to identify or distinguish infectious
CC strains of organisms in diagnostic procedures or in rickettsial
CC research where identification of particular organisms or strains is
CC needed.
XX
SQ Sequence 280 AA;

Query Match 81.8%; Score 90; DB 22; Length 280;
Best Local Similarity 88.9%; Pred. No. 2.3e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNTTGVFLKQDWGATI 18
   ||| ||||| ||||| |||
DB 60 nttagvfglkqdwgstl 77

RESULT 6
AAU06942
ID AAU06942 standard; Protein; 256 AA.
XX
AC AAU06942;
XX
DT 05-JUL-1999 (first entry)
XX
DE E. chaffeensis p28 protein.
XX
KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
KW detection; dog.
XX
OS Ehrlichia chaffeensis.
XX
PN WO9913720-A1.
XX
PD 25-MAR-1999.
XX
PF 18-SEP-1998; 98WO-US19600.

```


CC Anaplasma sp. or Cowdria sp. The vaccine elicits an immune response
 CC protective against the rickettsial pathogen. The nucleic acid vaccines
 CC can be driven by the human cytomegalovirus (HCMV) enhancer-promoter.
 CC Cowdria ruminantium genes designated map 2, lhworf3, 4hworf1, 18hworf1
 CC and 3gdrorf3 may be used in therapeutic and diagnostic applications. The
 CC polypeptides are useful for detecting antibodies associated with
 CC infection by a rickettsial pathogen whilst the polynucleotides may be
 CC used to detect the presence of rickettsial nucleic acids.

XX Sequence 276 AA;

Query Match 74.5%; Score 82; DB 21; Length 276;

Best Local Similarity 70.0%; Pred. No. 4.3e-05;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 NTTTGVFGLKQDWDGATIKD 20

||| |||||:||||:|:

Db 59 nttvgvfglkqwdgsaisn 78

RESULT 9

AAU04199
 ID AAU04199 standard; Protein: 276 AA.

XX AC AAU04199;

XX DT 23-OCT-2001 (first entry)

XX DE Variable surface antigen 5 (VSA5) from Ehrlichia chaffeensis.

XX KW Major antigenic protein; MAP; vaccine; immunogenic; rickettsia;
 infection; heartwater; diagnostic; variable surface antigen; VSA.

XX OS Ehrlichia chaffeensis.

XX PN US6251872-B1.

XX PD 26-JUN-2001.

XX PF 17-OCT-1997; 97US-0953326.

XX PR 17-OCT-1996; 96US-0733230.

XX PA (UYEL) UNIV FLORIDA.

XX PI Barbet AF, Ganta RR, McGuire TC, Burridge MJ, Nyika A;

PI Rurangirwa FR, Mahan SM, Bowie MV, Allenman AR;

XX DR WPI; 2001-424487/45.

XX DR N-PSDB; AAS07578.

XX PT New MAP2 genes and polypeptides useful as vaccines for conferring
 PT immunity to human and animal rickettsial diseases, e.g. heartwater, or
 PT as molecular markers in nucleic acid analysis procedures

XX PS Example 3; Fig 2A-2B; 30pp; English.

XX CC The sequence represents the amino acid sequence of variable surface
 CC antigen 5 (VSA5) isolated from Ehrlichia chaffeensis, which
 CC has similarity to major antigen protein (MAP). The MAP polynucleotides
 CC and polypeptides are useful as vaccines for conferring immunity to
 CC rickettsia infection, including Cowdria ruminantium causing heartwater.
 CC The MAP polynucleotides may be used as molecular markers in nucleic acid
 CC analysis procedures, and to produce the MAP polypeptides, which may
 CC be used to raise antibodies that are reactive with the polypeptides.
 CC The nucleic acids may further be used as probes to identify
 CC complementary sequences within other nucleic acid molecules or genomes,
 CC where such probes can be applied to identify or distinguish infectious
 CC strains of organisms in diagnostic procedures or in rickettsial
 CC research where identification of particular organisms or strains is
 CC needed.

SQ Sequence 276 AA;

Query Match 74.5%; Score 82; DB 22; Length 276;

Best Local Similarity 70.0%; Pred. No. 4.3e-05;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 NTTTGVFGLKQDWDGATIKD 20

||| |||||:||||:|:

Db 59 nttvgvfglkqwdgsaisn 78

RESULT 10

AAW51089

ID AAW51089 standard; Protein: 280 AA.

XX AC AAW51089;

XX DT 14-SEP-1998 (first entry)

XX DE Ehrlichia chaffeensis major antigenic protein 1 (MAP1).

XX KW MAP1 gene; major antigenic protein 1; rickettsia; DNA vaccine.

XX OS Ehrlichia chaffeensis.

XX PN WO9816554-A1.

XX PD 23-APR-1998.

XX PF 17-OCT-1997; 97WO-US19044.

XX PR 17-OCT-1996; 96US-0733230.

XX PA (UYEL) UNIV FLORIDA.

XX PI Barbet AF, Burridge MJ, Ganta RR, Mahan SM, McGuire TC;

PI Nyika A, Rurangirwa FR;

XX DR WPI; 1998-251232/22.

XX DR N-PSDB; AAV07177.

XX PT Composition containing nucleic acid encoding rickettsial antigen -
 PT useful for, e.g. stimulating protective immune response in humans or
 PT animals

XX PS Claim 3; Page 18-19; 39pp; English.

XX CC This polypeptide comprises the major antigen protein 1 gene (MAP1)
 CC of Ehrlichia chaffeensis. It is encoded by the MAP1 gene (see
 CC AAV07177). A claimed composition comprises a nucleic acid (see
 CC AAV07176-82) encoding a polypeptide (see AAW51088-99) that elicits a
 CC protective immune response against a rickettsial pathogen. The
 CC nucleic acid is used, in human or veterinary medicine, in vaccines
 CC to protect against Rickettsia, Ehrlichia, Anaplasma and Cowdria
 CC species. The nucleic acid does not replicate in the host but
 CC remains episomal and capable of expressing polypeptide for at least
 CC 19 mth. The Ehrlichia antigenic polypeptides can be used
 CC diagnostically to detect antibodies associated with Ehrlichia
 CC infection (claimed).

XX SQ Sequence 280 AA;

Query Match

Best Local Similarity 74.5%; Score 82; DB 19; Length 280;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 NTTTGVFGLKQDWDGATIKD 20

||| |||||:||||:|:

Db 60 nttvgvfglkqwdgsaisn 79

DE	Major antigenic protein 1 (MAP1) from Ehrlichia chaffeensis.
XX	Major antigenic protein 1; MAP1; vaccine; immunogenic; rickettsia;
XX	infection; heartwater; diagnostic.
XX	Ehrlichia chaffeensis.
OS	US6251872-B1.
PN	26-JUN-2001.
PPD	17-OCT-1997; 97US-0953326.
XX	17-OCT-1996; 96US-0733230.
XX	(UYFL) UNIV FLORIDA.
PA	Barbet AF, Ganta RR, McGuire TC, Burrige MJ, Nyika A;
XX	Rurangirwa FR, Mahan SM, Bowie MV, Alleman AR;
PPI	WPI: 2001-424487/45.
XX	N-PSDB; AAS07576.
DR	New MAP2 genes and polypeptides useful as vaccines for conferring
XX	immunity to human and animal rickettsial diseases, e.g. heartwater, or
PT	as molecular markers in nucleic acid analysis procedures -
PPT	Disclosure; Column 15-17; 30pp; English.
XX	The sequence represents the amino acid sequence of major antigenic
CC	protein 1 (MAP1) from Ehrlichia chaffeensis. The MAP polynucleotides and
CC	polypeptides are useful as vaccines for conferring immunity to rickettsia
CC	infection, including Cowdria ruminantium causing heartwater. The MAP
CC	polynucleotides may be used as molecular markers in nucleic acid
CC	analysis procedures, and to produce the MAP polypeptides, which may
CC	be used to raise antibodies that are reactive with the polypeptides.
CC	The nucleic acids may further be used as probes to identify
CC	complementary sequences within other nucleic acid molecules or genomes,
CC	where such probes can be applied to identify or distinguish infectious
CC	strains of organisms in diagnostic procedures or in rickettsial
CC	research where identification of particular organisms or strains is
CC	needed.
XX	Sequence 280 AA;
SQ	
Query Match	74.5%; Score 82; DB 22; Length 280;
Best Local Similarity	70.0%; Pred. No. 4.4e-05;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	
QY	1 NTTGVFGKQDWGATIKD 20 ::: :
Db	60 nttvgvfglkqndwgsaln 79
RESULT 13	
AAY06943	ID AAY06943 standard; Protein; 281 AA.
XX	AAY06943;
AC	05-JUL-1999 (first entry)
XX	E. chaffeensis OMP-1 protein.
DT	Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; p30;
XX	detection; dog.
KW	Ehrlichia chaffeensis.
XX	WO9913720-A1.
XX	25-MAR-1999.
PD	

XX 18-SEP-1998; 98WO-US19600.

XX WPI; 1999-254290/21.

XX N-PSDB; AAX34743.

XX (OHIS) UNIV OHIO STATE.

XX Ohashi N, Rikihisa Y;

XX WPI; 1999-254290/21.

XX N-PSDB; AAX34743.

XX Novel outer membrane proteins from Ehrlichia chaffeensis and

XX Ehrlichia canis

XX Disclosure; Fig 3B; 55pp; English.

XX The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.

XX Sequence 281 AA;

Query Match 74.5%; Score 82; DB 20; Length 281;

Best Local Similarity 70.0%; Pred. No. 4.4e-05;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTUTGVFGLKQDWDGATIKD 20

DB 59 nttvgvfglkqndwgsalsn 78

RESULT 14

AAY06962

ID AAY06962 standard; Protein; 280 AA.

AC AAY06962;

DT 05-JUL-1999 (first entry)

DE E. canis P30-2 protein.

XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;

XX detection; dog.

XX Ehrlichia canis.

XX WO9913720-A1.

XX 25-MAR-1999.

XX 18-SEP-1998; 98WO-US19600.

XX 19-SEP-1997; 97US-0059353.

XX (OHIS) UNIV OHIO STATE.

XX Ohashi N, Rikihisa Y;

XX WPI; 1999-254290/21.

XX N-PSDB; AAX34762.

XX Novel outer membrane proteins from Ehrlichia chaffeensis and

XX Ehrlichia canis

XX Disclosure; Fig 2B; 55pp; English.

XX The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part

CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.

XX Sequence 280 AA;

Query Match 70.0%; Score 77; DB 20; Length 280;

Best Local Similarity 65.0%; Pred. No. 0.00028;

Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NTUTGVFGLKQDWDGATIKD 20

DB 59 nstvgvfglkndwngttsn 78

RESULT 15

AAY71479

ID AAY71479 standard; Protein; 280 AA.

XX AAY71479;

DT 12-OCT-2000 (first entry)

XX Ehrlichia canis immunoreactive protein ECA28SA3.

XX Homologous 28-kDa protein gene; ECA28SA3; immunoreactive; vaccine;
 KW p28 gene; polymorphic multiple gene family; immunoprotective antigen;
 KW antibacterial; canine ehrlichiosis; canine tropical pancytopenia;
 KW tick-borne rickettsial disease; serodiagnosis.

XX Ehrlichia canis.

XX Key Location/Qualifiers

FT Peptide 1..23

FT /label= Signal_peptide

FT Protein 24..280

FT /label= Mature_ECA28SA3_28-kDa_protein

XX WO200032745-A2.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US28075.

XX 30-NOV-1998; 98US-0201458.

XX 03-MAR-1999; 99US-0261358.

XX (RERE-) RES DEV FOUND.

XX Walker DH, Yu X, McBride JW;

XX WPI; 2000-412298/35.

XX N-PSDB; AAD01294, AAD01295.

XX Ehrlichia canis antigens useful for vaccinating against canine

XX ehrlichiosis in dogs -

XX Claim 12; Page 68-69; 86pp; English.

XX The patent relates to homologous 28-kilodalton (kDa) protein genes of
 CC Ehrlichia canis, designated ECA28SA1, ECA28SA2, ECA28SA3, ECA28-1 and
 CC ECA28-2. These genes are members of a polymorphic multiple gene family
 CC and contained in a single locus of 5.592 kb. The 28-kDa proteins are
 CC immunoreactive with anti-E. canis serum hence are important
 CC immunoprotective antigens. The protein is useful for vaccinating
 CC against E. canis infections such as canine ehrlichiosis in dogs.
 CC Canine ehrlichiosis, also known as canine tropical pancytopenia, is a
 CC tick-borne rickettsial disease of dogs. ECA28-1 is conserved amongst
 CC different strains of E. canis and hence useful for serodiagnosis of
 CC canine ehrlichiosis. The present sequence is a E. canis
 CC ECA28SA3 30-kDa protein which is post-translationally modified to a

search completed:
Job time: 558 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2002, 14:27:08 ; Search time 13.08 seconds
(without alignments)
37.348 Million cell updates/sec

Title: US-10-054-647-2

Perfect score: 110

Sequence: 1 NTITGVFLKQDWDGATIKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
 - 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
 - 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
 - 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
 - 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
 - 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	81.8	280	4	US-08-953-326-17
2	82	74.5	276	4	US-08-953-326-18
3	82	74.5	280	3	US-08-733-230-4
4	82	74.5	280	4	US-08-953-326-4
5	71	64.5	286	4	US-08-953-326-15
6	59.5	54.1	287	3	US-08-733-230-2
7	59.5	54.1	287	4	US-08-953-326-2
8	59	53.6	278	4	US-08-953-326-16
9	49	44.5	133	4	US-08-953-326-20
10	40.5	36.8	1025	2	US-08-304-309-2
11	40.5	36.8	1025	3	US-08-991-942-2
12	40.5	36.8	1025	4	US-09-138-103-2
13	40.5	36.8	1025	5	PCT-US95-04567-4
14	40.5	36.8	2048	4	US-09-268-347-48
15	40	36.4	523	4	US-09-550-338-2
16	40	36.4	1212	4	US-09-268-866-2
17	39	35.5	182	3	US-08-828-741B-2
18	39	35.5	182	4	US-09-160-567-2
19	39	35.5	1025	2	US-08-304-309-4
20	39	35.5	1025	3	US-08-991-942-4
21	39	35.5	1025	5	PCT-US95-04567-2
22	38.5	35.0	569	2	US-08-467-822-27
23	38.5	35.0	569	4	US-08-432-697-27
24	38.5	35.0	569	4	US-08-466-248-27
25	38.5	35.0	859	1	US-08-053-614-2
26	38.5	35.0	859	1	US-08-316-397B-2
27	38.5	35.0	859	2	US-09-034-306-2

28	38.5	35.0	859	4	US-09-259-437-2	Sequence 2, Appl
29	38.5	35.0	859	5	PCT-US93-09782-2	Sequence 2, Appl
30	38.5	35.0	1181	1	US-08-053-614-4	Sequence 4, Appl
31	38.5	35.0	1181	1	US-08-316-397B-4	Sequence 4, Appl
32	38.5	35.0	1181	2	US-09-034-306-4	Sequence 4, Appl
33	38.5	35.0	1181	4	US-09-259-437-4	Sequence 4, Appl
34	38.5	35.0	1181	5	PCT-US93-09782-4	Sequence 4, Appl
35	38	34.5	135	1	US-08-446-600A-4	Sequence 4, Appl
36	38	34.5	247	4	US-09-199-637A-363	Sequence 363, App
37	38	34.5	338	4	US-09-413-231-6	Sequence 6, Appl
38	38	34.5	338	4	US-09-413-231-7	Sequence 7, Appl
39	38	34.5	371	3	US-08-586-165-3	Sequence 3, Appl
40	38	34.5	372	3	US-08-586-165-5	Sequence 5, Appl
41	38	34.5	548	1	US-08-333-358-2	Sequence 2, Appl
42	38	34.5	548	1	US-08-463-694-2	Sequence 2, Appl
43	38	34.5	548	1	US-08-694-501-2	Sequence 2, Appl
44	37.5	34.1	536	4	US-08-426-509A-12	Sequence 12, Appl
45	37.5	34.1	536	5	PCT-US95-05008-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-953-326-17
; Sequence 17, Application US/089533326
; Patent No. 6251872
; GENERAL INFORMATION:
; APPLICANT: Barbet, Anthony F.
; APPLICANT: Ganta, Roman R.
; APPLICANT: McGuire, Travis C.
; APPLICANT: Burridge, Michael J.
; APPLICANT: Nyika, Aceme
; APPLICANT: Rurangirwa, Fred R.
; APPLICANT: Mahan, Suman M.
; TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
; FILE REFERENCE: UP-167C1
; CURRENT APPLICATION NUMBER: US/08/953,326
; CURRENT FILING DATE: 1997-10-17
; EARLIER APPLICATION NUMBER: 08/953,326
; EARLIER FILING DATE: 1997-10-17
; EARLIER APPLICATION NUMBER: 08/733,230
; EARLIER FILING DATE: 1996-10-17
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
US-08-953-326-17

Query Match 81.8%; Score 90; DB 4; Length 280;
Best Local Similarity 88.9%; Pred. No. 4.9e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTITGVFLKQDWDGATIKD 18
||| ||||| ||||| |||
DB 60 NTITGVFLKQDWDGATIKD 77

RESULT 2
US-08-953-326-18
; Sequence 18, Application US/089533326
; Patent No. 6251872
; GENERAL INFORMATION:
; APPLICANT: Barbet, Anthony F.
; APPLICANT: Ganta, Roman R.
; APPLICANT: McGuire, Travis C.
; APPLICANT: Burridge, Michael J.
; APPLICANT: Nyika, Aceme
; APPLICANT: Rurangirwa, Fred R.

APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
FILE REFERENCE: UF-167C1
CURRENT APPLICATION NUMBER: US/08/953,326
CURRENT FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/953,326
EARLIER FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/733,230
EARLIER FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 18
LENGTH: 276
TYPE: PRT
ORGANISM: Ehrlichia chaffeensis
US-08-953-326-18

Query Match 74.5%; Score 82; DB 4; Length 276;
Best Local Similarity 70.0%; Pred. No. 9.5e-06;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NNTGVFGLKQDWDGATIKD 20
DB 59 NNTGVFGLKQDWDGSAISN 78

RESULT 3
US-08-733-230-4
Sequence 4, Application US/08/733230
Patent No. 6025338
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Ganta, Roman Reddy
APPLICANT: McGuire, Travis C.
APPLICANT: Burridge, Michael J.
APPLICANT: Nyika, Aceme
APPLICANT: Rurangirwa, Fred R.
APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against
TITLE OF INVENTION: Rickettsial Diseases and Methods of Use
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/733,230
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF-167
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-733-230-4

Query Match 74.5%; Score 82; DB 3; Length 280;
Best Local Similarity 70.0%; Pred. No. 9.6e-06;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NNTGVFGLKQDWDGATIKD 20
DB 60 NNTGVFGLKQDWDGSAISN 79

RESULT 4
US-08-953-326-4
Sequence 4, Application US/08953326
Patent No. 6251872
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Ganta, Roman R.
APPLICANT: McGuire, Travis C.
APPLICANT: Burridge, Michael J.
APPLICANT: Nyika, Aceme
APPLICANT: Rurangirwa, Fred R.
APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
TITLE OF INVENTION: Animals and Humans
FILE REFERENCE: UF-167C1
CURRENT APPLICATION NUMBER: US/08/953,326
CURRENT FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/953,326
EARLIER FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/733,230
EARLIER FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 280
TYPE: PRT
ORGANISM: Ehrlichia chaffeensis
US-08-953-326-4

Query Match 74.5%; Score 82; DB 4; Length 280;
Best Local Similarity 70.0%; Pred. No. 9.6e-06;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NNTGVFGLKQDWDGATIKD 20
DB 60 NNTGVFGLKQDWDGSAISN 79

RESULT 5
US-08-953-326-15
Sequence 15, Application US/08953326
Patent No. 6251872
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Ganta, Roman R.
APPLICANT: McGuire, Travis C.
APPLICANT: Burridge, Michael J.
APPLICANT: Nyika, Aceme
APPLICANT: Rurangirwa, Fred R.
APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
TITLE OF INVENTION: Animals and Humans
FILE REFERENCE: UF-167C1
CURRENT APPLICATION NUMBER: US/08/953,326
CURRENT FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/953,326
EARLIER FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/733,230
EARLIER FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 15
LENGTH: 286
TYPE: PRT
ORGANISM: Ehrlichia chaffeensis
US-08-953-326-15

Query Match 64.5%; Score 71; DB 4; Length 286;
Best Local Similarity 86.7%; Pred. No. 0.00059;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NTTGVFGLKQDWDGAI 18
||| ||||:|||||
Db 60 NTTGVFGLKQDWDRCVI 77

RESULT 6
US-08-733-230-2
Sequence 2, Application US/08733230
Patent No. 6025338
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Ganta, Roman R.
APPLICANT: McGuire, Travis C.
APPLICANT: Burridge, Michael J.
APPLICANT: Nyika, Aceme
APPLICANT: Rurangirwa, Fred R.
APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/733,230
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF-167
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 287 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-733-230-2

Query Match 54.1%; Score 59.5; DB 3; Length 287;
Best Local Similarity 80.0%; Pred. No. 0.043;
Matches 12; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 NTTGVFGLKQDWDG 15
||| ||||:|||||
Db 62 NTQT-VFGLKQDWDG 75

RESULT 7
US-08-733-230-2
Sequence 2, Application US/08733230
Patent No. 6025338
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Ganta, Roman R.
APPLICANT: McGuire, Travis C.
APPLICANT: Burridge, Michael J.
APPLICANT: Nyika, Aceme
APPLICANT: Rurangirwa, Fred R.
APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/733,230
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF-167
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 287 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-733-230-2

Query Match 54.1%; Score 59.5; DB 4; Length 287;
Best Local Similarity 80.0%; Pred. No. 0.043;
Matches 12; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 NTTGVFGLKQDWDG 15
||| ||||:|||||
Db 62 NTQT-VFGLKQDWDG 75

RESULT 8
US-08-953-326-16
Sequence 16, Application US/08953326
Patent No. 6251872
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Ganta, Roman R.
APPLICANT: McGuire, Travis C.
APPLICANT: Burridge, Michael J.
APPLICANT: Nyika, Aceme
APPLICANT: Rurangirwa, Fred R.
APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
FILE REFERENCE: UF-167C1
CURRENT APPLICATION NUMBER: US/08/953,326
CURRENT FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/953,326
EARLIER FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/733,230
EARLIER FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 278
TYPE: PRT
ORGANISM: Ehrlichia chaffeensis
US-08-953-326-16

Query Match 53.6%; Score 59; DB 4; Length 278;
Best Local Similarity 60.0%; Pred. No. 0.05;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTTGVFGLKQDWDG 15

US-08-953-326-2
Sequence 2, Application US/08953326
Patent No. 6251872
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Ganta, Roman R.
APPLICANT: McGuire, Travis C.
APPLICANT: Burridge, Michael J.
APPLICANT: Nyika, Aceme
APPLICANT: Rurangirwa, Fred R.
APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
FILE REFERENCE: UF-167C1
CURRENT APPLICATION NUMBER: US/08/953,326
CURRENT FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/953,326
EARLIER FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/733,230
EARLIER FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 287
TYPE: PRT
ORGANISM: Cowdria ruminantium
US-08-953-326-2

Query Match 54.1%; Score 59.5; DB 4; Length 287;
Best Local Similarity 80.0%; Pred. No. 0.043;
Matches 12; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 NTTGVFGLKQDWDG 15
||| ||||:|||||
Db 62 NTQT-VFGLKQDWDG 75

RESULT 8
US-08-953-326-16
Sequence 16, Application US/08953326
Patent No. 6251872
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Ganta, Roman R.
APPLICANT: McGuire, Travis C.
APPLICANT: Burridge, Michael J.
APPLICANT: Nyika, Aceme
APPLICANT: Rurangirwa, Fred R.
APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
FILE REFERENCE: UF-167C1
CURRENT APPLICATION NUMBER: US/08/953,326
CURRENT FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/953,326
EARLIER FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/733,230
EARLIER FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 278
TYPE: PRT
ORGANISM: Ehrlichia chaffeensis
US-08-953-326-16

Query Match 53.6%; Score 59; DB 4; Length 278;
Best Local Similarity 60.0%; Pred. No. 0.05;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTTGVFGLKQDWDG 15

```
Db 60 NPTVYGLKQDWEG 74
| | :|:|||||:|
| | :|:|||||:|

RESULT 9
US-08-953-326-20
; Sequence 20, Application US/08953326
; Patent No. 6251872
; GENERAL INFORMATION:
; APPLICANT: Barbet, Anthony F.
; APPLICANT: Ganta, Roman R.
; APPLICANT: McGuire, Travis C.
; APPLICANT: Burridge, Michael J.
; APPLICANT: Nyika, Aceme
; APPLICANT: Rurangirwa, Fred R.
; APPLICANT: Mahan, Suman M.
; TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases of
; FILE REFERENCE: UP-167C1
; CURRENT APPLICATION NUMBER: US/08/953,326
; CURRENT FILING DATE: 1997-10-17
; EARLIER APPLICATION NUMBER: 08/953,326
; EARLIER FILING DATE: 1997-10-17
; EARLIER APPLICATION NUMBER: 08/733,230
; EARLIER FILING DATE: 1996-10-17
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 20
; LENGTH: 133
; TYPE: prt
; ORGANISM: Ehrlichia canis
US-08-953-326-20

Query Match 44.5%; Score 49; DB 4; Length 133;
Best Local Similarity 56.2%; Pred. No. 0.89;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 TTGVFGLKQDWGATI 18
| | :|:|||||:|
| | :|:|||||:|
Db 62 TTVYGLKENWAGDAI 77

RESULT 10
US-08-304-309-2
; Sequence 2, Application US/08304309
; Patent No. 5856454
; GENERAL INFORMATION:
; APPLICANT: GONZALEZ, Frank J.
; APPLICANT: FERNANDEZ-SALGUERO, Pedro
; TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 09-SEP-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US/08/304,309
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-991-942-2

Query Match 36.8%; Score 40.5; DB 2; Length 1025;
Best Local Similarity 42.9%; Pred. No. 2.2e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 1 NTTGVFGLKQD---WDGATI 18
| | :|:|||||:|
| | :|:|||||:|
Db 736 NTVSGLMGLKSDGTPWPAVGI 756

RESULT 11
US-08-991-942-2
; Sequence 2, Application US/08991942
; Patent No. 6015673
; GENERAL INFORMATION:
; APPLICANT: GONZALEZ, Frank J.
; APPLICANT: FERNANDEZ-SALGUERO, Pedro
; TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 09-SEP-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US/08/304,309
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-991-942-2

Query Match 36.8%; Score 40.5; DB 3; Length 1025;
Best Local Similarity 42.9%; Pred. No. 2.2e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 1 NTTGVFGLKQD---WDGATI 18
| | :|:|||||:|
| | :|:|||||:|
Db 736 NTVSGLMGLKSDGTPWPAVGI 756
```

```

; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04567-4

Query Match          36.88;   Score 40.5;   DB 5;   Length 1025;
Best Local Similarity 42.9%;   Pred. No. 2.2e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY      1  NNTTGVFLKQD---WDGATI 18
      || :|| ||| | | |
Db      736  NTVSGLMGLASDGTTPAVGI 756

RESULT 14
US-09-268-347-48
; Sequence 48, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 2048
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-48

```

```

Query Match          36.8%; Score 40.5; DB 4; Length 2048;
Best Local Similarity 39.3%; Pred. No. 5e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 11; Gaps 1;

QY      - 4 TGVPGLKQDW-----DGATIKD 20
      |  | | | |
DB      686 TVTFTGLSQDSGLTIGKSTLNDGLTVKD 713

RESULT 15
US-09-550-338-2
; Sequence 2, Application US/09550338
; Patent No. 6210951
; GENERAL INFORMATION:
; APPLICANT: KAWASAKI, Hisashi
; APPLICANT: USUDA, Yoshihiro
; APPLICANT: MIHARA, Yoshihiro
; APPLICANT: KURAHASHI, Osamu
; TITLE OF INVENTION: GMP Synthetase and Gene Coding for the Same
; FILE REFERENCE: 0010-1101-0
; CURRENT APPLICATION NUMBER: US/09/550,338
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: JP 11-114787
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum (Corneybacterium glutamicum)
US-09-550-338-2

Query Match          36.4%; Score 40; DB 4; Length 523;
Best Local Similarity 41.2%; Pred. No. 1.2e+02;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OV      4 TGVPGLKQDWDGATIKD 20

```

Db 190 TEIAGLEQNTAANIAE 206

Search completed: June 7, 2002, 14:30:36
Job time: 206 sec

GenCore version 4.5
Copyright (c) 1993-2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2002, 14:27:58 ; Search time 15.01 Seconds
(without alignments)
128.034 Million cell updates/sec

Title: US-10-054-647-2

Perfect score: 110
Sequence: 1 NTTVGVFLKQDWDGATIKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 71.*

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	81.8	280	JE0217	28k surface antige
2	82	74.5	276	JE0218	28k surface antige
3	71	64.5	286	JE0219	28k surface antige
4	59	53.6	278	JE0216	28k surface antige
5	59	53.6	284	I40882	major antigenic pr
6	52	47.3	5188	B85547	probable RTX fami
7	52	47.3	5291	F90896	hypothetical prote
8	51	46.4	1191	A53491	bumetanide-sensiti
9	49	44.5	133	JE0221	28k surface antige
10	49	44.5	634	F97172	flagellar hook-ass
11	47	42.7	540	S54586	probable membrane
12	47	42.7	584	C48858	flagellin - Escher
13	46.5	42.3	287	C83342	conserved hypothet
14	46	41.8	261	G84057	hypothetical prote
15	45	40.9	160	A75466	2-demethylmenaquin
16	45	40.9	756	T20109	hypothetical prote
17	44	40.0	281	AH3012	outer surface prot
18	44	40.0	284	B69345	phage-related prot
19	44	40.0	284	G98271	hypothetical prote
20	44	40.0	465	S47738	cytochrome-c perox
21	44	40.0	465	F91178	cytochrome-c perox
22	44	40.0	465	G86024	cytochrome-c perox
23	44	40.0	534	C82096	aminoacyl-histidin
24	44	40.0	648	P3BFF6	F3 protein - phage
25	44	40.0	1004	JH0470	Na+/K+-exchanging
26	43.5	39.5	290	S76787	hypothetical prote
27	43	39.1	280	D70976	hypothetical prote
28	43	39.1	437	I40167	dihydroorotase (EC
29	43	39.1	470	JC4098	tetracycline 6-hyd

30 39.1 482 2 G75483 probable leucyl am
31 43 681 2 G87276 hypothetical prote
32 43 1037 2 A56594 Na+/K+-exchanging
33 43 1649 2 C86822 hypothetical prote
34 43 1944 2 AH3098 rhizobioicin/RTX to
35 43 1990 2 A96188 probable phosphoes
36 42.5 123 2 AE0293 conserved hypothet
37 42 160 2 A12476 hypothetical prote
38 42 224 2 C72390 hypothetical prote
39 42 287 1 S56603 probable formate a
40 42 287 2 A91296 probable activatin
41 42 287 2 D86137 probable activatin
42 42 307 2 T44893 hypothetical prote
43 42 362 2 F75379 S-adenosylmethioni
44 42 413 2 T08297 conserved hypothet
45 42 656 2 A41870 dnaA protein - Str

ALIGNMENTS

RESULT 1
JE0217
28k surface antigen 4 - Ehrlichia chaffensis
N:Alternate names: MAP1
C:Species: Ehrlichia chaffensis
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 17-Mar-1999
C:Accession: JE0217
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tri
A:Reference number: JE0216; MUID:98321180
A:Accession: JE0217
A:Molecule type: DNA
A:Residues: 1-280 <RED>
A:Cross-references: GB:AF062761

Query Match 81.8%; Score 90; DB 2; Length 280;
Best Local Similarity 88.9%; Pred. No. 1.8e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTTVGVFLKQDWDGATI 18
Db 60 NTTVGVFLKQDWDGSTI 77
|||||

RESULT 2
JE0218
28k surface antigen 5 - Ehrlichia chaffensis
N:Alternate names: MAP1
C:Species: Ehrlichia chaffensis
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 17-Mar-1999
C:Accession: JE0218
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tri
A:Reference number: JE0216; MUID:98321180
A:Accession: JE0218
A:Molecule type: DNA
A:Residues: 1-276 <RED>
A:Cross-references: GB:AF062761

Query Match 74.5%; Score 82; DB 2; Length 276;
Best Local Similarity 70.0%; Pred. No. 3.1e-05;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTTVGVFLKQDWDGATIKD 20
Db 59 NTTVGVFLKQDWDGSAIN 78
|||||

RESULT 3
 JEO219
 28k surface antigen 2 - Ehrlichia chaffensis
 N:Alternate names: MAP1
 C:Species: Ehrlichia chaffensis
 C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 17-Mar-1999
 C:Accession: JEO219
 R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burrige, M.J.; Alleman, A.R.
 Biochem. Biophys. Res. Commun. 247, 636-643, 1998
 A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe
 A:Reference number: JEO216; MUID:98321180
 A:Accession: JEO219
 A:Molecule type: DNA
 A:Residues: 1-286 <RED>
 A:Cross-references: GB:AF062761

Query Match 64.5%; Score 71; DB 2; Length 286;
 Best Local Similarity 66.7%; Pred. No. 0.0017;
 Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NTTGVFGLKQDWDGAT 18
 ||| ||||:|||||
 Db : 60 NTTGVFGLKQDWDRCVY 77

RESULT 4
 JEO216
 28k surface antigen 3 - Ehrlichia chaffensis
 N:Alternate names: MAP1
 C:Species: Ehrlichia chaffensis
 C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 17-Mar-1999
 C:Accession: JEO216
 R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burrige, M.J.; Alleman, A.R.
 Biochem. Biophys. Res. Commun. 247, 636-643, 1998
 A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe
 A:Reference number: JEO216; MUID:98321180
 A:Accession: JEO216
 A:Molecule type: DNA
 A:Residues: 1-278 <RED>
 A:Cross-references: GB:AF062761

Query Match 53.6%; Score 59; DB 2; Length 278;
 Best Local Similarity 60.0%; Pred. No. 0.12;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTTGVFGLKQDWDG 15
 . | | : ||||| |
 Db 60 NPTVALYGLKQDWDG 74

RESULT 5
 I40882
 major antigenic protein - heartwater rickettsia
 C:Species: Cowdria ruminantium (heartwater rickettsia)
 C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 08-Oct-1999
 C:Accession: I40882; S42827
 R:van Vliet, A.H.; Jongejan, F.; van Kleef, M.; van der Zeijst, B.A.
 Infect. Immun. 62, 1451-1456, 1994
 A:Title: Molecular cloning, sequence analysis, and expression of the gene encoding the
 A:Reference number: I40882; MUID:94178956
 A:Accession: I40882
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-284 <RES>
 A:Cross-references: EMBL:X74250; NID:G454266; PIDN:CAA52309.1; PID:G454267
 C:Genetics:
 A:Gene: map1

Query Match 53.6%; Score 59; DB 2; Length 284;
 Best Local Similarity 76.9%; Pred. No. 0.13;

Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TTGVFGLKQDWDG 15
 | ||||| ||||
 Db 63 TKAVFGLKQDWDG 75

RESULT 6
 B85547
 probable RTX family exoprotein [imported] - Escherichia coli (strain O157:H7, substra
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: B85547
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: B85480; MUID:21074935; PMID:11206551
 A:Accession: B85547
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-5188 <STO>
 A:Cross-references: GB:AE005174; NID:G12513368; PIDN:AAG54838.1; GSPDB:GN00145; UWGP:
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: Z0615

Query Match 47.3%; Score 52; DB 2; Length 5188;
 Best Local Similarity 60.0%; Pred. No. 37;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 TTGVFGLKQDWDGA 16
 ||||| : |||||
 Db 4803 TTSGVAAMDYDWDGA 4817

RESULT 7
 F90696
 hypothetical protein ECs0542 [imported] - Escherichia coli (strain O157:H7, substrain
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: F90696
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: F90696
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-5291 <HAV>
 A:Cross-references: GB:BA000007; PIDN:BA033965.1; PID:G13360000; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509932
 C:Genetics:
 A:Gene: ECs0542

Query Match 47.3%; Score 52; DB 2; Length 5291;
 Best Local Similarity 60.0%; Pred. No. 38;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 TTGVFGLKQDWDGA 16
 ||||| : |||||
 Db 4906 TTSGVAAMDYDWDGA 4920

RESULT 8
 A53491
 bumetanide-sensitive Na-K-Cl cotransporter - spiny dogfish
 C:Species: Squalus acanthias (spiny dogfish)
 C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 02-Mar-2001
 C:Accession: A53491

```

Query Match          44.5%   Score: 49;   DB 2;   Length 634;
Best Local Similarity 50.0%   Pred. No. 11;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 NNTTGVFLKQDWDGATIKD 20
      :| |||| :| | ||| :|

```

QY 1 NNTTGVFLKOD 12
 |||||:||||:
 Db 199 NNTTGLYGLATE 210
 RESULT 13

C83242
Conserved hypothetical protein PA3239 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83242
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, S.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337
A:Accession: C83242
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <STO>
A:Cross-references: GB:AE004746; GB:AE004091; NID:g9949350; PIDN:AAG06627.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3239

Query Match 42.3%; Score 46.5; DB 2; Length 267;
Best Local Similarity 64.7%; Pred. No. 11;
Matches 11; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 1 NTGTGFGKQDWDGAT 17
|||||
DB 145 NTGTGVLGL---WDPAT 158

RESULT 14

G84057
hypothetical protein BH3263 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G84057
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai, N.
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G84057
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-261 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB06982.1; GSPDB:GN001
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3263
C:Superfamily: Bacillus subtilis hypothetical protein ytmp

Query Match 41.8%; Score 46; DB 2; Length 261;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 DWGATIKD 20
|||||
DB 177 DWGATVAD 185

RESULT 15

A75466
2-demethylmenaquinone 2-C-methyltransferase (EC 2.1.1.-) DR0859 [similarity] - Deinococcus
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C:Accession: A75466
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M., M.; Shen, M.; Vamathevan, J.O.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Ma, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: A75466

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-160 <WHI>
A:Cross-references: GB:AE001940; GB:AE000513; NID:g6458577; PIDN:AAF10437.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0859
A:Map position: 1
C:Keywords: methyltransferase
Query Match 40.9%; Score 45; DB 2; Length 160;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 5 GVFGKQDWDGATI 18
|||||
DB 80 GVFGVNGWEGVII 93

Search completed: June 7, 2002, 14:30:58
Job time: 180 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2002, 14:30:38 Search time 10.34 Seconds
(without alignments)
74.893 Million cell updates/sec

Title: US-10-054-647-2

Perfect score: 110

Sequence: 1 NITTVGLKQWDGATKND 20

Scoring table: BLOSUM62

Gapop 10.0 Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	46.4	1191	1	NCCL_SQAC
2	47	42.7	540	1	Y8M_YEAST
3	45	40.9	160	1	MENG_DEIRA
4	45	40.9	756	1	K6PF_CABEL
5	44	40.0	284	1	YQAK_BACSU
6	44	40.0	465	1	YHJA_ECOLI
7	44	40.0	648	1	VP3_BPRH6
8	44	40.0	1004	1	AT1A_ARTSF
9	43.5	39.5	230	1	AROE_SYNY3
10	43	39.1	427	1	PYRC_BACCL
11	43	39.1	1021	1	ALAI_CANFA
12	42	38.2	287	1	XJW_ECOLI
13	42	38.2	307	1	YM16_MYLE
14	42	38.2	656	1	DNAI_SFRCO
15	42	38.2	1020	1	ALAI_HUMAN
16	41.5	37.7	366	1	GCST_NEIMB
17	41.5	37.7	368	1	GCST_NEIMA
18	41	37.3	377	1	YA67_METH
19	41	37.3	395	1	KIME_RAT
20	41	37.3	684	1	HTPG_PORGI
21	41	37.3	749	1	VPA_ROTGA
22	41	37.3	866	1	YCBS_ECOLI
23	40.5	36.8	196	1	ANFL_CHICK
24	40.5	36.8	734	1	PURL_ZYMO
25	40.5	36.8	1025	1	DPYD_HUMAN
26	40	36.4	84	1	GVM1_HALN1
27	40	36.4	120	1	YRAJ_BACSU
28	40	36.4	191	1	PGHD_URSAR
29	40	36.4	282	1	BI08_HELPJ
30	40	36.4	282	1	BI08_HELPJ
31	40	36.4	341	1	X33B_MYCPN
32	40	36.4	489	1	UBPT_CABEL
33	40	36.4	500	1	AMPA_BACSU

34 40 36.4 942 1 ENV_CAEVG
35 40 36.4 1020 1 AIA2_RAT
36 40 36.4 1023 1 AIA1_HUMAN
37 40 36.4 1041 1 ATNA_DROME
38 40 36.4 1205 1 NKCL_MOUSE
39 40 36.4 1212 1 NKCL_HUMAN
40 40 36.4 1295 1 GLPI_CAEEL
41 40 36.4 1429 1 LIL2_CAEEL
42 39 35.5 105 1 YGRM_MICEC
43 39 35.5 118 1 ANFD_RANCA
44 39 35.5 149 1 DTD_CLOAB
45 39 35.5 267 1 NUSL_ASPOR

ALIGNMENTS

RESULT 1
NKCL_SQAC
ID NKCL_SQAC STANDARD; PRT: 1191 AA.
AC P55013;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Bimetanide-sensitive sodium-(Potassium)-chloride cotransporter 1
DE (NA-K-CL symporter) (NKCC).
GN SLC12A2 OR NKCC1.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE FROM N.A., AND PHOSPHORYLATION SITES.
RC TISSUE=Rectal gland;
RX MEDLINE=94181560; PubMed=8134373;
RA Xu J.-C., Lytle C., Zhu T.T., Payne J.A., Benz E. Jr., Forbush B. III;
RT "Molecular cloning and functional expression of the
bimetanide-sensitive Na-K-Cl cotransporter."
RL Proc. Natl. Acad. Sci. U.S.A. 91:2201-2205(1994).
CC -I- FUNCTION: ELECTRICALLY SILENT TRANSPORTER SYSTEM WHICH IS A
MEDIATOR OF SODIUM AND CHLORIDE REABSORPTION. PLAYS A VITAL ROLE
IN THE REGULATION OF IONIC BALANCE AND CELL VOLUME.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- TISSUE SPECIFICITY: EXPRESSED IN MANY TISSUES.
CC -I- SIMILARITY: BELONGS TO THE SLC12A FAMILY OF TRANSPORTERS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; U05958; AAB60617.1;
DR InterPro; IPR002293; AA_rel_permease_1.
DR InterPro; IPR002027; Amino_acid_permease.
DR InterPro; IPR002443; NaKCl_transporter.
DR Pfam; PF00324; aa_permeases; 1.
DR PRINTS; PR01207; NAKCLTRNSPRT.
KW Transport; Transmembrane; Glycoprotein; Phosphorylation.
FT DOMAIN 1 257 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 258 278 POTENTIAL.
FT DOMAIN 279 281 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 282 302 POTENTIAL.
FT DOMAIN 303 339 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 340 360 POTENTIAL.
FT DOMAIN 361 382 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 383 403 POTENTIAL.
FT DOMAIN 404 407 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 408 428 POTENTIAL.
FT DOMAIN 429 458 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 459


```
QY 5 GVFGKQDWGATIK 18
DB 80 GVFGVNGWEGVII 93

RESULT 4
KQPF_CAEEL STANDARD; PRT; 756 AA.
AC KQPF_CAEEL
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable 6-phosphofructokinase (EC 2.7.1.11) (Phosphofructokinase)
DE (Phosphohexokinase)
GN C50F4.2
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL-N2;
RA McMurray A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate -> ADP + D-
fructose 1,6-bisphosphate.
CC -1- PATHWAY: KEY CONTROL STEP OF GLYCOLYSIS.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOFRUCTOKINASE FAMILY, TWO DOMAIN
SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; Z70750; CAA94737.1;
DR HSSP; P00512; 3PFK.
DR WormPep; C50F4.2; CE05467.
DR InterPro; IPR000023; Phosphofructokinase.
DR Pfam; PF00365; PFK; 2.
DR PRINTS; PR00476; PFRCKINASE.
DR ProDom; P000707; Phosphofructokinase; 2.
DR ProSITE; PS00433; PHOSPHOFRUCTOKINASE; 1.
DR Kinase; Transferase; Glycolysis; Repeat.
SQ SEQUENCE 756 AA; 83301 MW; 26A89B801D286534 CRC64;

Query Match 40.9%; Score 45; DB 1; Length 756;
Best Local Similarity 53.3%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 6 VFGKQDWGATIKD 20
DB 436 VIGIKHGDGLKND 450

RESULT 5
YQAK_BACSU STANDARD; PRT; 284 AA.
AC YQAK_BACSU
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 32.2 kDa protein in SPOIIC-CMLA intergenic region.
GN YQAK
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]

SEQUENCE FROM N.A.
STRAIN=168 / JH642;
MEDLINE=95219086; PubMed=7704261;
Takamaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;
"Complete nucleotide sequence of a skin element excised by DNA
rearrangement during sporulation in Bacillus subtilis.";
Microbiology 141:323-327(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RA Kobayashi Y., Mizuno M., Masuda S., Takamaru K., Hosono S., Sato T.,
Takeuchi M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP IDENTIFICATION.
MEDLINE=96084975; PubMed=7489895;
Medigue C., Moszer I., Vliar A., Danchin A.;
"Analysis of a Bacillus subtilis genome fragment using a co-operative
computer system prototype.";
Gene 165:GC37-GC51(1995).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; D32216; BRA06925.1;
DR EMBL; D84432; BRA12386.1;
DR EMBL; Z99117; CAB14569.1;
DR Subtilist; BG11262; Yqak.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 284 AA; 32170 MW; F255261D4692ADB7 CRC64;

Query Match 40.0%; Score 44; DB 1; Length 284;
Best Local Similarity 53.8%; Pred. No. 6.7;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 FGLKQDWGATIK 19
DB 193 FGKNDWDAMALK 205

RESULT 6
YHJA_ECOLI STANDARD; PRT; 465 AA.
AC YHJA_ECOLI
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable cytochrome C peroxidase (EC 1.1.1.5).
GN YHJA OR B3518.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
"Analysis of the Escherichia coli genome. V. DNA sequence of the
region from 76.0 to 81.5 minutes.";
Nucleic Acids Res. 22:2576-2586(1994).
CC -1- CATALYTIC ACTIVITY: 2 ferrocyclochrome c + 2 H(2)O.
ferrocyclochrome c + 2 H(2)O.
CC -1- PTM: BINDS 3 HEMES (POTENTIAL).
CC -1- SIMILARITY: HIGH, TO P.AERUGINOSA CYTOCHROME C551 PEROXIDASE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U00039; AAB18494.1;
 CC EMBL; AEO00428; AAC76543.1;
 CC EcoGene; EG12244; YhJA.
 CC InterPro; IPR000345; CytC_heme_bind.
 CC InterPro; IPR003088; Cyt_C1.
 CC Pfam; PF00034; cytochrome_c; 1.
 CC PROSITE; PS00190; CYTOCHROME_C; 3.
 CC Hypothetical protein; Oxidoreductase; Peroxidase; Heme;
 CC Electron transport; Complete proteome.
 CC BINDING 59 59 HEME 1 (COVALENT) (BY SIMILARITY).
 CC BINDING 62 62 HEME 1 (COVALENT) (BY SIMILARITY).
 CC METAL 63 63 IRON 1 (HEME PROXIMAL) (BY SIMILARITY).
 CC BINDING 207 207 HEME 2 (COVALENT) (BY SIMILARITY).
 CC BINDING 210 210 HEME 2 (COVALENT) (BY SIMILARITY).
 CC METAL 211 211 IRON 2 (HEME PROXIMAL) (BY SIMILARITY).
 CC BINDING 351 351 HEME 3 (COVALENT) (BY SIMILARITY).
 CC BINDING 354 354 HEME 3 (COVALENT) (BY SIMILARITY).
 CC METAL 355 355 IRON 3 (HEME PROXIMAL) (BY SIMILARITY).
 CC METAL 429 429 IRON 3 (HEME DISTAL) (BY SIMILARITY).
 CC SEQUENCE 465 AA; 51570 MW; 9F49AA698949E6DA CRC64;

Query Match 40.08; Score 44; DB 1; Length 465;
 Best Local Similarity 52.99; Pred. No. 11;
 Matches 9; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

OY 6 VFGKQDWDG--ATIKD-20
 DB 243 VFNEQFDGRLQD 259
 ||:|:|:|:|:|:|:|:|:|

RESULT 7
 VP3_BPPH6 STANDARD; PRT; 648 AA.
 AC P11129;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 01-JAN-1990 (Rel. 13, Last annotation update)
 DE P3 protein.
 GN P3
 OS Bacteriophage phi-6.
 CC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
 CC NCBI_TaxID=10879;
 RN [1].
 CC SEQUENCE FROM N.A., AND SEQUENCE OF 1-13.
 RX MEDLINE=88160044; PubMed=3347997;
 RA Gottlieb P., Metzger S., Romantschuk M., Carton J., Strassman J.,
 RA Bamford D.H., Kalkkainen N., Mindich L.;
 RT "Nucleotide sequence of the middle dRNA segment of bacteriophage phi
 6: placement of the genes of membrane-associated proteins";
 RL Virology 163:183-190(1988).
 CC -1- FUNCTION: P3 PROTEIN IS NECESSARY FOR ADSORPTION ONTO HOST
 CC CELLS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M17462; AAA68485.1;
 CC PIR; C28648; P3BPP6.
 CC Envelope protein
 CC SEQUENCE 648 AA; 69178 MW; B188DFE02ACC54E3 CRC64;

Query Match 40.08; Score 44; DB 1; Length 648;
 Best Local Similarity 41.2%; Pred. No. 16;
 Matches 7; Conservative 5; Mismatches 3; Indels 2; Gaps 1;

OY 6 VFG--LKQDWDGATIKD 20
 DB 181 IFGWYKMDWEGSAVAD 197
 ||:|:|:|:|:|:|:|:|:|

RESULT 8
 ATIA_ARTSF
 ID ATIA_ARTSF STANDARD; PRT; 1004 AA.
 AC P28774;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sodium/potassium-transporting ATPase alpha chain (EC 3.6.3.9) (Sodium
 pump) (Na+/K+ ATPase).
 DE Artemia sanfranciscana (Brine shrimp) (Artemia franciscana).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
 OC Artemiidae; Artemia.
 OC NCBI_TaxID=6661;
 RN [1].
 CC SEQUENCE FROM N.A.
 CC MEDLINE=92039032; PubMed=1657719;
 CC Macias M.I., Martinez J.L., Palmero I., Sastre L.;
 CC "Cloning of a cDNA encoding an Artemia franciscana Na/K ATPase alpha-
 CC subunit";
 CC Gene 105:197-204(1991).
 CC -1- FUNCTION: THIS IS THE CATALYTIC COMPONENT OF THE ACTIVE ENZYME,
 CC WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE EXCHANGE OF
 CC NA AND K IONS ACROSS THE PLASMA MEMBRANE. THIS ACTION CREATES THE
 CC ELECTROCHEMICAL GRADIENT OF NA AND K, PROVIDING THE ENERGY FOR
 CC ACTIVE TRANSPORT OF VARIOUS NUTRIENTS.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+)(In) + K(+)(Out) = ADP +
 CC phosphate + Na(+)(Out) + K(+)(In).
 CC -1- SUBUNIT: COMPOSED OF THREE SUBUNITS: ALPHA (CATALYTIC), BETA
 CC AND GAMMA.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IIC.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; X56650; CAA39972.1;
 CC PIR; JH0470; JH0470.
 CC HSSP; P04191; 1EUL.
 CC InterPro; IPR004014; Cation_ATPase.
 CC InterPro; IPR001757; E1-E2_ATPase.
 CC InterPro; IPR001454; Hydrolase.
 CC InterPro; IPR000661; Na_K_ATPase.
 CC Pfam; PF00690; Cation_ATPase_C; 1.
 CC Pfam; PF00690; Cation_ATPase_N; 1.
 CC Pfam; PF00122; E1-E2_ATPase; 1.
 CC Pfam; PF00702; Hydrolase; 1.
 CC PRINTS; PR00119; CATATPASE.
 CC PRINTS; PR00121; NAKATPASE.
 CC PROSITE; PS00154; ATPASE_E1_E2; 1.
 CC Hydrolase; Sodium/potassium transport; Transmembrane;
 CC Phosphorylation; ATP-binding.
 CC TRANSMEM 76 96 BY SIMILARITY.
 CC TRANSMEM 110 126 BY SIMILARITY.
 CC TRANSMEM 272 294 BY SIMILARITY.
 CC TRANSMEM 301 329 BY SIMILARITY.
 CC TRANSMEM 768 791 BY SIMILARITY.

FT TRANSHEM 828 855 BY SIMILARITY.
 FT TRANSHEM 897 918 BY SIMILARITY.
 FT TRANSHEM 934 959 PHOSPHORYLATION (PROBABLE).
 FT MOD_RES 357 357
 FT BINDING 489 489 ATP (BY SIMILARITY).
 SQ SEQUENCE 1004 AA; 110699 MW; CB4E6BCE19A78C7 CRC64;

Query Match 40.0%; Score 44; DB 1; Length 1004;
 Best Local Similarity 40.0%; Pred. No. 26;
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 6 VFGKQDWDGATIKD 20
 DB 864 LFGURKWDGRAVND 878

RESULT 9
 AROE_SYNY3
 ID AROE_SYNY3 STANDARD; PRT; 290 AA.
 AC P74591;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Shikimate 5-dehydrogenase (EC 1.1.1.25).
 GN AROE OR SLR159.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97081201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugiyama M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions."
 RL DNA Res. 3:109-136(1996).
 CC -1- CATALYTIC ACTIVITY: Shikimate + NADP(+) -> 5-dehydroshikimate +
 NADPH.
 CC -1- PATHWAY: FOURTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE
 BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -1- SIMILARITY: BELONGS TO THE SHIKIMATE DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch)
 CC
 CC EMBL; D90916; BAA18699.1;
 CC InterPro; IPR002907; Shikimate_DH.
 CC Pfam; PF01488; Shikimate_DH.1.
 CC Aromatic amino acid biosynthesis; Oxidoreductase; NADP;
 CC Complete proteome.
 SQ SEQUENCE 290 AA; 31099 MW; 8A2D38EE5D57B303 CRC64;

Query Match 39.5%; Score 43.5; DB 1; Length 290;
 Best Local Similarity 47.8%; Pred. No. 8.3;
 Matches 11; Conservative 0; Mismatches 5; Indels 7; Gaps 1;

OY 2 TTTCVFG-----LKQDWDGAT 17
 DB 105 TNDTVEGFLAPLPLEKQDWSGRT 127

RESULT 10

PYRC_BACCL
 ID PYRC_BACCL STANDARD; PRT; 427 AA.
 AC P46538;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Dihydroorotase (EC 3.5.2.3) (DHOase).
 GN PYRC.
 OS Bacillus caldolyticus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=DSM 405;
 RX MEDLINE=94282293; PubMed=7516791;
 RA Ghim S.Y., Nielsen P., Neuhaud J.,
 RT "Molecular characterization of pyrimidine biosynthesis genes from the
 thermophile Bacillus caldolyticus."
 RL Microbiology 140:479-491(1994).
 CC -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O -> N-carbamoyl-L-
 aspartate.
 CC -1- COFACTOR: THIS ENZYME TIGHTLY BINDS ONE ZINC ATOM PER CHAIN WHICH
 IS REQUIRED FOR THE CATALYTIC MECHANISM. IT ALSO BINDS WEAKLY TO
 TWO OTHERS ZINCS WHICH ARE NOT ESSENTIAL FOR ACTIVITY
 (BY SIMILARITY).
 CC -1- PATHWAY: THIRD STEP IN PYRIMIDINE BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE DHOASE FAMILY. SUBFAMILY 2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch)
 CC
 CC EMBL; X73308; CAA51737.1;
 CC MEROPS; M38.972;
 CC InterPro; IPR002195; Dihydroorotase.
 CC Pfam; PF00744; Dihydroorotase.1.
 CC PROSITE; PS00483; DIHYDROOROTASE_1; 1.
 CC PROSITE; PS00483; DIHYDROOROTASE_2; 1.
 CC Pyrimidine biosynthesis; Hydrolyase; Zinc.
 FT METAL 60 62 ZINC (POTENTIAL).
 FT METAL 62 62 ZINC (POTENTIAL).
 SQ SEQUENCE 427 AA; 46047 MW; 759A2AA99F733F4E CRC64;

Query Match 39.1%; Score 43; DB 1; Length 427;
 Best Local Similarity 68.8%; Pred. No. 15;
 Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 TGVEGLKQDWDGATIK 19
 DB 342 TGVFTLKQLVLDLTIK 357

RESULT 11
 ALAL_CANFA
 ID ALAL_CANFA STANDARD; PRT; 1021 AA.
 AC P50997;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sodium/potassium-transporting, ATPase alpha-1 chain precursor
 (EC 3.6.3.9) (Sodium pump 1) (Na+/K+ ATPase 1).
 GN ATP1A1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;

```

TRANSMEM 984 1004
DOMAIN 1005 1021
MOD_RES 16 16
MOD_RES 374 374
MOD_RES 941 941
BINDING 80 82
METAL 715 715
METAL 719 719
CONFLICT 109 109
CONFLICT 185 185
CONFLICT 249 249
CONFLICT 309 309
SEQUENCE 1021 AA; 112666 MW; 938A19AA487CBEAA CRC64;

Query Match 39.1%; Score 43; DB 1; Length 1021;
Best Local Similarity 47.1%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

4 TGVFGLKQDWGATIKD 20
879 THLLGLRVDWDRIND 895

RESULT 12
ACID VJWV ECOLI STANDARD; PRT; 287 AA.
P39409;
01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical protein.yjjw.
VJWV OR B4379.
Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
NCBI_TaxId=562;
SEQUENCE FROM N.A.
STRAIN=K12 / MG1655;
MEDLINE=95334362; PubMed=7610040;
Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
Blattner F.R.;
"Analysis of the Escherichia coli genome VI: DNA sequence of the
region from 92.8 through 100 minutes."
Nucleic Acids Res. 23:2105-2119(1995).
-|- CORACTOR: BINDS 1 4FE-4S CLUSTER (BY SIMILARITY).
-|- SIMILARITY: BELONGS TO THE ORGANIC RADICAL ACTIVATING ENZYMES
FAMILY.
-|- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/annouce/
or send an email to license@isb-sib.ch).
-----
EMBL; U14003; AAA97275.1;
EMBL; AE000508; RAC77332.1;
HSP; P00196; 2FDN.
Ecogene; EG12599; YJJW.
InterPro; IPR001450; 4Fe4S_ferredoxin.
InterPro; IPR001989; Radical_activ.
Pfam; PF00037; fer4; 2;
Pfam; PF02143; Radical_activat; 1.
ProDom; PD004758; Radical_activat; 1.
PROSITE; PS00198; 4FE4S-FERREDOXIN; 2.
PROSITE; PS01087; RADICAL_ACTIVATING; 1.

```

```

QY 9 LQQDWDGATI 18
| | | | |
DDB 143 LQDWDGATL 152

RESULT 14
DDBNA_STRCO STANDARD; PRT; 656 AA;
P27902; O9KX4;
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotator update)
Chromosomal replication initiator protein dnaA.
DNA OR, SCH18.16C.
Streptomyces coelicolor.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
NCBI_TaxID-1902;
[1]
SEQUENCE FROM N.A.
STRAIN-A3(2);
MEDLINE-92250416; PubMed-1577691;
Calcutt M.J., Schmidt F.J.;
"Conserved gene arrangement in the origin region of the Streptomycetes
coelicolor chromosome.";
J. Bacteriol. 174:3220-3226(1992).
[2]
SEQUENCE OF 51-656 FROM N.A.
STRAIN-A3(2);
Brown S.P., Harris D., Cerdeno A.M., Parkhill J., Barrell B.G.,
Rajandream M.A.;
Submitted (MAY 2000) to the EMBL/GenBank/DDJF databases.
! FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
(DNA BOX): 5'-TTATC(C/A)(C/A)-3'. DNAA BINDS TO ATP AND TO
ACIDIC PHOSPHOLIPIDS.
! SIMILARITY: BELONGS TO THE DNAA FAMILY.
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use, by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to: license@isb-sib.ch).
-----
EMBL; AF187159; AAA26734.1;
EMBL; AL357152; CAB92999.1;
PIR; A41870; A41870.
InterPro; IPR003593; AAA.
InterPro; IPR001957; Bac.DnaA.
Pfam; PF00308; bac_dnaa; 1.
PRINTS; PR00051; DNAA.
SMART; SM00382; AAA; 1.
PROSITE; PS01008; DNAA; 1.
DNA replication; DNA-binding; ATP-binding.
NP_BIND 357 364 ATP (POTENTIAL).
SEQUENCE 656 AA; 73182 MW; 6CID5C0193D3C92B CRC64;

Query Match 38.2%; Score 42; DB 1; Length 656;
Best Local Similarity 36.8%; Pred. No. 35;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 TTGTGVLKQDWDGATIKD 20
| | | | |
DDB 563 TSTAINGATADYFGLTVED 581

RESULT 15
ID A1A2 HUMAN STANDARD; PRT; 1020 AA.

```

AC PS0993; Q07059;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-OCT-2001 (Rel. 40, Last annotation update)
 DE Sodium/potassium-transporting ATPase alpha-2 chain precursor
 DE (EC 3.6.3.9) (Sodium pump 2) (Na+/K+ ATPase 2).
 GN ATP1A2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90008924; PubMed=2477373;
 RX Shull M.M., Pugh D.G., Lingrel J.B.;
 RT "Characterization of the human Na,K-ATPase alpha 2 gene and
 RT identification of intragenic restriction fragment length
 RT polymorphisms.";
 RL J. Biol. Chem. 264:17532-17543(1989).
 RN [2]
 RP SEQUENCE OF 211-249 FROM N.A.
 RC TISSUE=Leukocyte;
 RX MEDLINE=87231946; PubMed=3035563;
 RA Shull M.M., Lingrel J.B.;
 RT "Multiple genes encode the human Na+,K+-ATPase catalytic subunit.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4039-4043(1987).
 RN [3]
 RP SEQUENCE OF 251-442 FROM N.A.
 RC TISSUE=Placenta, and Brain;
 RX MEDLINE=87247232; PubMed=3036582;
 RA Sverdlov E.D., Monastyrskaya G.S., Broude N.E., Ushkaryov Y.A.,
 RA Allikmets R.L., Melkov A.M., Smirnov Y.V., Malyshov I.V.,
 RA Kostlova I.E., Petrukhin K.E., Grishin A.V., Kijatkin N.I.,
 RA Kustova M.B., Sverdlov V.E., Modyanov N.N., Ovchinnikov Y.A.;
 RT "The family of human Na+ K+-ATPase genes: No less than five genes
 RT and/or pseudogenes related to the alpha-subunit.";
 RL FEBS Lett. 217:275-278(1987).
 RN [4]
 RP SEQUENCE OF 1-4 FROM N.A.
 RX MEDLINE=89153603; PubMed=2537767;
 RA Sverdlov E.D., Bessarab D.A., Malyshov I.V., Petrukhin K.E.,
 RA Smirnov Y.V., Ushkaryov Y.A., Monastyrskaya G.S., Broude N.E.,
 RA Modyanov N.N.;
 RT "Family of human Na+,K+-ATPase genes. Structure of the putative
 RT regulatory region of the alpha+-gene.";
 RL FEBS Lett. 244:481-483(1989).
 CC -1- FUNCTION: THIS IS THE CATALYTIC COMPONENT OF THE ACTIVE ENZYME,
 CC NA AND K IONS ACROSS THE HYDROLYSIS OF ATP COUPLED WITH THE EXCHANGE OF
 CC ELECTROCHEMICAL GRADIENT OF NA AND K, PROVIDING THE ENERGY FOR
 CC ACTIVE TRANSPORT OF VARIOUS NUTRIENTS.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+)(In) + K(+)(Out) = ADP +
 CC phosphate + Na(+)(Out) + K(+)(In).
 CC -1- SUBUNIT: COMPOSED OF THREE SUBUNITS: ALPHA (CATALYTIC), BETA
 CC AND GAMMA.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IIC.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J05096; AAA51797.1;
 CC EMBL; M16795; AAA51799.1;
 CC EMBL; M27578; AAA35575.1;
 CC EMBL; M27571; AAA35575.1; JOINED.
 CC EMBL; M27576; AAA35575.1; JOINED.
 CC EMBL; Y07494; CAA68793.1; ALT_SEQ.

DR HSP; P04191; LEUL.
 DR MIM; 182340;
 DR InterPro; IPR004014; Cation_ATPase.
 DR InterPro; IPR001757; E1-E2_ATPase.
 DR InterPro; IPR001454; Hydrolase.
 DR InterPro; IPR000661; Na_K_ATPase.
 DR Pfam; PF00889; Cation_ATPase_C; 1.
 DR Pfam; PF00690; Cation_ATPase_N; 1.
 DR Pfam; PF00122; E1-E2_ATPase; 1.
 DR Pfam; PF00702; Hydrolase; 1.
 DR PRINTS; PR00119; CATATPASE.
 DR PRINTS; PR00121; NAKATPASE.
 DR PROSITE; PS00154; ATPASE_E1_E2; 1.
 DR KW Hydrolase; Sodium/potassium transport; Transmembrane; Phosphorylation;
 KW Magnesium; Metal-binding; ATP-binding; Multigene family.
 FT PROPEP 1 5
 FT CHAIN 6 1020
 FT DOMAIN 6 85
 FT TRANSMEM 86 106
 FT DOMAIN 107 129
 FT TRANSMEM 130 150
 FT DOMAIN 151 286
 FT TRANSMEM 287 306
 FT DOMAIN 307 318
 FT TRANSMEM 319 336
 FT DOMAIN 337 769
 FT TRANSMEM 770 789
 FT DOMAIN 790 799
 FT TRANSMEM 800 820
 FT DOMAIN 821 840
 FT TRANSMEM 841 863
 FT DOMAIN 864 916
 FT TRANSMEM 916 935
 FT DOMAIN 936 948
 FT TRANSMEM 949 967
 FT DOMAIN 968 982
 FT TRANSMEM 983 1003
 FT DOMAIN 1004 1020
 FT MOD_RES 374 374
 FT MOD_RES 940 940
 FT BINDING 80 82
 FT METAL 714 714
 FT METAL 718 718
 FT SEQUENCE 1020 AA; 112265 MW; AFBD8EA94FFB4FC3 CRC64;
 SQ
 Query Match 38.2%; Score 42; DB 1; Length 1020;
 Best Local Similarity 40.0%; Pred. No. 57;
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 6 VFGLKQDWDGATIKD 20
 Db 880 LIGRLDWDRTMND 894

Search completed: June 7, 2002, 14:33:58
 Job time: 200 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2002, 14:30:18 ; Search time 25.04 seconds
(without alignments)
138.175 Million cell updates/sec

Title: US-10-054-647-2

Perfect score: 110

Sequence: 1 NTTCVFLKQWDGATIKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum: DB seq length: 0

Maximum: DB seq length: 2000000000

Post-processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Database : SPREMBL19:

- 1: sp_archaea:
- 2: sp_bacteria:
- 3: sp_fungi:
- 4: sp_human:
- 5: sp_invertebrate:
- 6: sp_mammal:
- 7: sp_mhc:
- 8: sp_organelle:
- 9: sp_phage:
- 10: sp_plant:
- 11: sp_rodent:
- 12: sp_virus:
- 13: sp_vertebrate:
- 14: sp_unclassified:
- 15: sp_rvirus:
- 16: sp_bacteriap:
- 17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	288	Q9ZGJ2	Q9ZGJ2 ehrlichia c
2	96	87.3	280	Q52107	Q52107 ehrlichia c
3	90	81.8	280	Q85357	Q85357 ehrlichia c
4	86	78.2	246	Q9RH35	Q9RH35 ehrlichia c
5	86	78.2	275	Q93DD4	Q93DD4 ehrlichia c
6	86	78.2	276	Q85817	Q85817 ehrlichia c
7	86	78.2	276	Q93DD1	Q93DD1 ehrlichia c
8	82	74.5	276	Q85358	Q85358 ehrlichia c
9	82	74.5	281	Q9AC19	Q9AC19 ehrlichia c
10	82	74.5	281	Q93DD2	Q93DD2 ehrlichia c
11	81	73.6	280	Q9ZGM9	Q9ZGM9 ehrlichia c
12	81	73.6	280	Q85816	Q85816 ehrlichia c
13	81	73.6	280	Q93DD3	Q93DD3 ehrlichia c
14	77	70.0	280	Q9ADV3	Q9ADV3 ehrlichia c
15	77	70.0	280	Q9F473	Q9F473 ehrlichia c
16	76	69.1	276	Q9F475	Q9F475 ehrlichia c

17	71	64.5	286	2	Q52105	Q52105 ehrlichia c
18	70	63.6	278	2	Q9F472	Q9F472 ehrlichia c
19	70	63.6	278	2	Q9R8A9	Q9R8A9 ehrlichia c
20	70	63.6	278	2	Q9R8A8	Q9R8A8 ehrlichia c
21	70	63.6	278	2	Q9R8A7	Q9R8A7 ehrlichia c
22	70	63.6	278	2	Q9R8A6	Q9R8A6 ehrlichia c
23	70	63.6	278	2	Q9R8A5	Q9R8A5 ehrlichia c
24	70	63.6	278	2	Q9R3J3	Q9R3J3 ehrlichia c
25	70	63.6	307	2	Q9ZGJ1	Q9ZGJ1 ehrlichia c
26	61.5	55.9	270	2	Q9AF98	Q9AF98 cowdria rum
27	61.5	55.9	277	2	Q93E57	Q93E57 cowdria rum
28	61.5	55.9	278	2	Q93E56	Q93E56 cowdria rum
29	61.5	55.9	278	2	Q93E52	Q93E52 cowdria rum
30	61.5	55.9	290	2	Q46324	Q46324 cowdria rum
31	61.5	55.9	290	2	Q46332	Q46332 cowdria rum
32	61.5	55.9	290	2	Q93E64	Q93E64 cowdria rum
33	61.5	55.9	290	2	Q93E63	Q93E63 cowdria rum
34	60	54.5	290	2	Q9AEU3	Q9AEU3 cowdria rum
35	59.5	54.1	265	2	Q9AF99	Q9AF99 cowdria rum
36	59.5	54.1	275	2	Q93E59	Q93E59 cowdria rum
37	59.5	54.1	276	2	Q93E60	Q93E60 cowdria rum
38	59.5	54.1	276	2	Q93E53	Q93E53 cowdria rum
39	59.5	54.1	287	2	Q9R425	Q9R425 cowdria rum
40	59.5	54.1	287	2	Q46329	Q46329 cowdria rum
41	59.5	54.1	287	2	Q46331	Q46331 cowdria rum
42	59.5	54.1	290	2	Q46333	Q46333 cowdria rum
43	59.5	54.1	290	2	Q46330	Q46330 cowdria rum
44	59	53.6	272	2	Q9AMF6	Q9AMF6 cowdria sp.
45	59	53.6	272	2	Q93E54	Q93E54 cowdria rum

ALIGNMENTS

RESULT 1

ID	Q9ZGJ2	PRELIMINARY;	PRT;	288 AA.
AC	Q9ZGJ2;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DE	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	30-KDA MAJOR OUTER MEMBRANE PROTEIN (P28-8).			
GN	P30 OR P28-8.			
OS	Ehrlichia canis.			
OC	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;			
OC	Rickettsiaceae; Ehrlichieae; Ehrlichia.			
OX	NCBI_TaxID:944;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-OKLAHOMA;			
RA	MEDLINE-98371112; PubMed-9705412;			
RA	Ohashi N., Unver A., Zhi N., Rikihisa Y.;			
RT	"Cloning and characterization of multigenes encoding the			
RT	immunodominant 30-kilodalton major outer membrane proteins of			
RT	Ehrlichia canis and application of the recombinant protein for			
RT	serodiagnosis."			
RL	J. Clin. Microbiol. 36:2671-2680(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-JAKE;			
RA	MEDLINE-99242757; PubMed-10225842;			
RA	McBride J.W., Yu Xj, Walker D.H.;			
RT	"Molecular cloning of the gene for a conserved major immunoreactive			
RT	28-kilodalton protein of Ehrlichia canis: a potential serodiagnostic			
RT	antigen."			
RL	Clin. Diagn. Lab. Immunol. 6:392-399(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-JAKE;			
RA	MEDLINE-20432107; PubMed-10974556;			
RA	McBride J.W., Yu X.J., Walker D.H.;			
RT	"A conserved, transcriptionally active p28 multigene locus of			
RT	Ehrlichia canis."			

5

AC	Q9AC19;				
DT	01-JUN-2001	(TRENBLrel. 17, Created)			
DT	01-JUN-2001	(TRENBLrel. 17, Last sequence update)			
DT	01-DEC-2001	(TRENBLrel. 19, Last annotation update)			

DE MAJOR OUTER MEMBRANE PROTEIN P28.
 OS Ehrlichia chaffeensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichieae; Ehrlichia.
 OX NCBI_TaxID=945;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-ARKANSAS;
 RC MEDLINE=98084465; PubMed=9423849;
 RX Choshi N., Zhi N., Zhang Y., Rikihisa Y.
 RA "Immunodominant major outer membrane proteins of Ehrlichia chaffeensis
 RT are encoded by a polymorphic multigene family."
 RL Infect. Immun. 66:132-139(1998).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-ARKANSAS;
 RC MEDLINE=21153566; PubMed=11254561;
 RX Ohashi N., Rikihisa Y., Unver A.
 RA "Analysis of transcriptionally Active Gene Clusters of Major Outer
 RT Membrane Protein Multigene Family in Ehrlichia canis and E.
 RT chaffeensis."
 RL Infect. Immun. 69:2083-2091(2001).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN-V1;
 RC Long S.W., Zhang X.-F., Qi H., Standaert S., Walker D.H., Yu X.-J.;
 RA "Allele variation and patterns of transcription of the Ehrlichia
 RT chaffeensis 28 kDa outer membrane protein multigene family."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U72291; AAC28673.1;
 DR EMBL: AF393388; AAL12918.1;
 DR InterPro: IPR002566; Surface_Ag_msp4.
 DR Pfam: PF01617; Surface_Ag_2_1;
 DR SEQUENCE 281 AA; 30343 MW; A995F7C4459AA9A CRC64;
 SQ

Query Match 74.5%; Score 82; DB 2; Length 281;
 Best Local Similarity 70.0%; Pred. No. 8.7e-05;
 Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTTGVFGLKQDWDGATIKD 20
 ||| |||||:||||:|
 Db 59 NTTGVFGLKQDWDGSAISN 78

RESULT 10
 ID Q93DD2 PRELIMINARY; PRT; 281 AA.
 AC Q93DD2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE OUTER MEMBRANE PROTEIN P28.
 OS Ehrlichia chaffeensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichieae; Ehrlichia.
 OX NCBI_TaxID=945;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-V6;
 RC Long S.W., Zhang X.-F., Qi H., Standaert S., Walker D.H., Yu X.-J.;
 RA "Allele variation and patterns of transcription of the Ehrlichia
 RT chaffeensis 28 kDa outer membrane protein multigene family."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF393392; AAL12922.1;
 DR EMBL: AF393392; AAL12922.1;
 SQ SEQUENCE 281 AA; 30388 MW; B54DBB55AC839A9A CRC64;

Query Match 74.5%; Score 82; DB 2; Length 281;
 Best Local Similarity 70.0%; Pred. No. 8.7e-05;
 Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTTGVFGLKQDWDGATIKD 20
 ||| |||||:||||:|
 Db 59 NTTGVFGLKQDWDGSAISN 78

RESULT 11
 ID Q9ZGM9 PRELIMINARY; PRT; 280 AA.
 AC Q9ZGM9;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE OUTER MEMBRANE PROTEIN P28.
 OS Ehrlichia chaffeensis
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichieae; Ehrlichia.
 OX NCBI_TaxID=945;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-JAX;
 RC MEDLINE=99175287; PubMed=10074538;
 RX Yu X.-J., McBride J.W., Walker D.H.;
 RA "Genetic diversity of the 28-kilodalton outer membrane protein gene in
 RT human isolates of Ehrlichia chaffeensis."
 RL J. Clin. Microbiol. 37:1137-1143(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-JAX;
 RC Yu X.-J., Walker D.H.;
 RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF077733; AAC31546.1;
 DR InterPro: IPR002566; Surface_Ag_msp4.
 DR Pfam: PF01617; Surface_Ag_2_1;
 DR SEQUENCE 280 AA; 30308 MW; 91C5AC7851B77F2 CRC64;
 SQ

Query Match 73.6%; Score 81; DB 2; Length 280;
 Best Local Similarity 77.8%; Pred. No. 0.00012;
 Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTTGVFGLKQDWDGATIKD 18
 ||| |||||:||||:|
 Db 59 STTAGVFLKQDWDGSAI 76

RESULT 12
 ID O85816 PRELIMINARY; PRT; 280 AA.
 AC O85816;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE OUTER MEMBRANE PROTEIN P28.
 OS Ehrlichia chaffeensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichieae; Ehrlichia.
 OX NCBI_TaxID=945;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-91HE17;
 RC MEDLINE=99175287; PubMed=10074538;
 RX Yu X.-J., McBride J.W., Walker D.H.;
 RA "Genetic diversity of the 28-kilodalton outer membrane protein gene in
 RT human isolates of Ehrlichia chaffeensis."
 RL J. Clin. Microbiol. 37:1137-1143(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-91HE17;
 RC Yu X.-J., Walker D.H.;
 RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF393392; AAL12922.1;
 DR EMBL: AF393392; AAL12922.1;
 SQ SEQUENCE 281 AA; 30388 MW; B54DBB55AC839A9A CRC64;

Query Match 74.5%; Score 82; DB 2; Length 281;
 Best Local Similarity 70.0%; Pred. No. 8.7e-05;
 Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

RA Long S.W., Zhang X.-F., Qi H., Standaert S., Walker D.H., Yu X.-J.;
RT "Allele variation and patterns of transcription of the Ehrlichia
RT chaffeensis 28 kDa outer membrane protein multigene family";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF077732; AAC31545.1; -
DR EMBL: AF393394; AAL12924.1; -
DR EMBL: AF393390; AAL12920.1; -
DR InterPro: IPR002566; Surface_Ag_msp4.
DR Pfam: PF01617; Surface_Ag_2; 1.
SQ SEQUENCE 280 AA; 30277 MW; 91C54AC78507A63F CRC64;

Query Match 73.6%; Score 81; DB 2; Length 280;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGTGVLKQDWDGATI 18
DB 59 STTAGVFLKQDWDGSAI 76

RESULT 13

ID Q93DD3 PRELIMINARY; PRT; 280 AA.
AC Q93DD3; 2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE OUTER MEMBRANE PROTEIN P28.
OS Ehrlichia chaffeensis
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Ehrlichieae; Ehrlichia.
OX NCBI_TaxID=945;
[1]
RN STRAIN-V5;
RP SEQUENCE FROM N.A.
RC
RA Long S.W., Zhang X.-F., Qi H., Standaert S., Walker D.H., Yu X.-J.;
RT "Allele variation and patterns of transcription of the Ehrlichia
RT chaffeensis 28 kDa outer membrane protein multigene family";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF393391; AAL12921.1; -
SQ SEQUENCE 280 AA; 30372 MW; C7B8C8710BC167E9 CRC64;

Query Match 73.6%; Score 81; DB 2; Length 280;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGTGVLKQDWDGATI 18
DB 59 STTAGVFLKQDWDGSAI 76

RESULT 14

ID Q9ADV3 PRELIMINARY; PRT; 280 AA.
AC Q9ADV3; 2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN P30-2.
OS Ehrlichia canis.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Ehrlichieae; Ehrlichia.
OX NCBI_TaxID=944;
[1]
RN STRAIN-OKLAHOMA;
RP SEQUENCE FROM N.A.
RC
RX MEDLINE=98371112; PubMed=9705412;
RA Ohashi N., Unver A., Zhi N., Rikhiya Y.;
RT "Cloning and characterization of multigenes encoding the
RT immunodominant 30-kilodalton major outer membrane proteins of

RT Ehrlichia canis and application of the recombinant protein for
RT serodiagnosis";
RL J. Clin. Microbiol. 36:2671-2680(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-OKLAHOMA;
RX MEDLINE=21153566; PubMed=11254561;
RA Ohashi N., Rikhiya Y., Unver A.;
RT "Analysis of Transcriptionally Active Gene Clusters of Major Outer
RT Membrane Protein Multigene Family in Ehrlichia canis and E.
RT chaffeensis";
RL Infect. Immun. 69:2083-2091(2001).
DR EMBL: AF078553; AAK28699.1; -
DR InterPro: IPR002566; Surface_Ag_msp4.
DR Pfam: PF01617; Surface_Ag_2; 1.
SQ SEQUENCE 280 AA; 30803 MW; 27238BE1C7E68A91 CRC64;

Query Match 70.0%; Score 77; DB 2; Length 280;
Best Local Similarity 65.0%; Pred. No. 0.00052;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NTGTGVLKQDWDGATI 20
DB 59 NSTGVFLKHDWNGGTISN 78

RESULT 15

ID Q9F473 PRELIMINARY; PRT; 280 AA.
AC Q9F473; 2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE P28-6.
GN P28-6.
OS Ehrlichia canis.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Ehrlichieae; Ehrlichia.
OX NCBI_TaxID=944;
[1]
RN STRAIN-JAKE;
RP SEQUENCE FROM N.A.
RC
RX MEDLINE=99242757; PubMed=10225842;
RA McBride J.W., Yu X.J., Walker D.H.;
RT "Molecular cloning of the gene for a conserved major immunoreactive
RT 28-kilodalton protein of Ehrlichia canis: a potential serodiagnostic
RT antigen";
RL Clin. Diagn. Lab. Immunol. 6:392-399(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC
RX MEDLINE=20432107; PubMed=10974556;
RA McBride J.W., Yu X.J., Walker D.H.;
RT "A conserved, transcriptionally active p28 multigene locus of
RT Ehrlichia canis";
RL Gene 254:245-252(2000).
DR EMBL: AF082744; AAG14361.1; -
DR InterPro: IPR002566; Surface_Ag_msp4.
DR Pfam: PF01617; Surface_Ag_2; 1.
SQ SEQUENCE 280 AA; 30762 MW; BE284A4B94FE3123 CRC64;

Query Match 70.0%; Score 77; DB 2; Length 280;
Best Local Similarity 65.0%; Pred. No. 0.00052;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NTGTGVLKQDWDGATI 20
DB 59 NSTGVFLKHDWNGGTISN 78

Search completed: June 7, 2002, 14:33:41
Job time: 203 sec

THIS PAGE BLANK (USPTO)